

Art Unit: 2655

Amendments to the Claims:

This listing of claims will replace all prior versions, and listings, of claims in the application:

5

Listing of Claims:

1-4. (canceled)

10 5. (original) A method of controlling high-speed reading in a text-to-speech conversion system including a text analysis module for generating a phoneme and prosody character string from an input text; a prosody generation module for generating a synthesis parameter of at least a voice segment, a phoneme duration, and a fundamental frequency for the phoneme and prosody character string;
15 a voice segment dictionary in which voice segments as a source of voice are registered; and a speech generation module for generating a synthetic waveform by waveform superimposition by referring to said voice segment dictionary,

20 said method comprising the step of providing said prosody generation module with a sound quality coefficient determination unit that has a sound quality conversion coefficient table for changing said voice segment to switch sound quality and selects from said sound quality conversion coefficient table such a coefficient that sound quality does not change when a user-
25 designated utterance speed exceeds a threshold.

6. (original) The method according to claim 5, wherein said threshold is a predetermined maximum utterance speed.

30 7. (previously presented) A method of controlling high-speed reading in a text-to-speech conversion system including a text

Art Unit: 2655

analysis module for generating a phoneme and prosody character string from an input text; a prosody generation module for generating a synthesis parameter of at least a voice segment, phoneme duration, and fundamental frequency for the phoneme and prosody character string; a voice segment dictionary in which voice segments as a source of voice are registered; and a speech generation module for generating a synthetic waveform by waveform superimposition by referring to said voice segment dictionary, said method comprising the step of providing said prosody generation module with both a pitch contour correction unit for outputting a pitch contour corrected according to an intonation level designated by the user and a switch for determining whether a base pitch is added to said pitch contour corrected according to said user-designated utterance speed, said switch being controlled not to change the base pitch when the utterance speed exceeds a threshold.

8. (original) The method according to claim 7, wherein said threshold is a predetermined maximum utterance speed.

20

9. (original) The method according to claim 7, wherein said pitch contour correction unit performs a pitch contour generation process that includes a phrase component calculation process in which all phrases of an input sentence are processed by calculating a phrase component by statistical analysis according to said user-designated utterance speed or making said phrase component zero and a process in which all words in said input sentence are processed by calculating an accent component by statistical analysis according to said user-designated utterance speed and either correcting said accent component according to said user designated intonation level or making said accent component zero.

30

Art Unit: 2655

10. (original) A method of controlling high-speed reading in a text-to-speech conversion system including a text analysis module for generating a phoneme and prosody character string from an input text; a prosody generation module for generating a synthesis parameter of at least a voice segment, a phoneme duration, and a fundamental frequency for said phoneme and prosody character string; a voice segment dictionary in which voice segments as a source of voice are registered; and a speech generation module for generating a synthetic waveform by waveform superimposition while referring to said voice segment dictionary,

said method comprising the step of providing said speech generation module with signal sound generation means for inserting a signal sound between sentences to indicate an end of a sentence when a user-designated utterance speed exceeds a threshold.

11. (original) The method according to claim 10, wherein said threshold is a predetermined maximum utterance speed.

12. (original) A method of controlling high-speed reading in a text-to-speech conversion system including a text analysis module for generating a phoneme and prosody character string from an input text; a prosody generation module for generating a synthesis parameter of at least a voice segment, a phoneme duration, and a fundamental frequency for the phoneme and prosody character string; a voice segment dictionary in which voice segments as a source of voice are registered; and a speech generation module for generating a synthetic waveform by waveform superimposition by referring to said voice segment dictionary,

said method comprising the step of providing said prosody generation module with a phoneme duration determination unit for

Art Unit: 2655

performing a process in which when a user-designated utterance speed exceeds a threshold, an utterance speed of at least a leading word in a sentence is returned to a normal utterance speed.

5 13. (original) The method according to claim 12, wherein said threshold is a predetermined maximum utterance speed.

10 14. (original) The method according to claim 12, wherein said phoneme duration determination unit performs a process in which when a word under process is a leading word in a sentence and said user-designated utterance speed exceeds said threshold, a phoneme duration is not corrected and, when said word under process is not a leading word of a sentence or said user-designated utterance speed does not exceed said threshold, a first process by which a
15 phoneme duration correction coefficient is changed according to said user-designated utterance speed and a second process in which all syllables of said word are processed by correcting a length of a vowel or vowels of said word, and carrying out said first and second processes for all words contained in the sentence.

20

15. (previously presented) A method of controlling high-speed reading in a text-to-speech conversion system, comprising:
inputting a text into the text-to-speech conversion system;
generating a phoneme and prosody character string of the text
25 with a text analysis module;
creating a duration rule table containing a first phoneme duration obtained empirically;
creating a duration prediction table containing a second phoneme duration obtained through statistical analysis;
30 designating an utterance speed;
determining a threshold value;

Art Unit: 2655

comparing the utterance speed with the threshold value;
selecting one of the duration rule table and the duration
prediction table according to the utterance speed;

determining a third phoneme duration with a phoneme duration
5 determination unit according to the one of the duration rule table
and the duration prediction table;

generating a synthesis parameter of at least a voice segment,
the third phoneme duration, and a fundamental frequency of the
phoneme and prosody character string with a prosody generation
10 module; and

generating a synthetic waveform through waveform
superimposition with a speech generation module according to the
synthesis parameter and a voice segment dictionary containing a
voice segment as a basic source of voice.

15

16. (previously presented) The method according to claim 15,
in the step of selecting the one of the duration rule table and the
duration prediction table according to the utterance speed, said
duration rule table is selected when the utterance speed exceeds
20 the threshold value, and said duration prediction table is selected
when the utterance speed does not exceed the threshold value.

17. (previously presented) The method according to claim 15,
in the step of determining the threshold value, said threshold
25 value is determined to be a predetermined maximum utterance speed.

18. (previously presented) A method of controlling high-speed
reading in a text-to-speech conversion system, comprising:
inputting a text into the text-to-speech conversion system;
30 generating a phoneme and prosody character string of the text
with a text analysis module;

Art Unit: 2655

creating a rule table containing first data of accent and phrase components obtained empirically;

creating a prediction table containing second data of accent and phrase components obtained through statistical analysis;

5 designating an utterance speed;

determining a threshold value;

comparing the utterance speed with the threshold value;

selecting one of the rule table and the prediction table according to the utterance speed;

10 determining a pitch contour with a pitch contour determination unit according to the one of the rule table and the prediction table;

generating a synthesis parameter of at least a voice segment, a phoneme duration, and a fundamental frequency of the phoneme and prosody character string with a prosody generation module; and

15 generating a synthetic waveform through waveform

superimposition with a speech generation module according to the synthesis parameter and a voice segment dictionary containing a voice segment as a basic source of voice.

20

19. (previously presented) The method according to claim 18, in the step of selecting the one of the rule table and the prediction table according to the utterance speed, said rule table is selected when the utterance speed exceeds the threshold value, and said prediction table is selected when the utterance speed does not exceed the threshold value.

25

20. (previously presented) The method according to claim 18; in the step of determining the threshold value, said threshold value is determined to be a predetermined maximum utterance speed.

30